



DADOS ABERTOS NA PESQUISA EM EDUCAÇÃO EM CIÊNCIAS: PERSPECTIVAS, DESAFIOS E POSSIBILIDADES

Nathália Helena Azevedo^I

<https://orcid.org/0000-0003-2075-9121>

Paula Cristina Cardoso Mendonça^{II}

<https://orcid.org/0000-0003-1762-4840>

RESUMO:

Neste editorial apresentamos as perspectivas de ciência aberta adotada pelo periódico dando destaque para as práticas de dados abertos e dados compartilhados. A disponibilização de dados de pesquisa pode aprimorar a transparência, a colaboração, a reprodutibilidade, a replicação, o reuso e o enriquecimento da produção científica no campo da Educação em Ciências. Por isso, neste editorial, trazemos novas perspectivas à temática com base em nossos aprendizados ao longo de 2023 na implementação de uma editoria de dados, a qual detalhamos o trabalho em um fluxograma. Além disso, abordamos algumas das possíveis implicações do compartilhamento ético de dados na área e relatamos alguns dos frutos das interações entre nossa equipe editorial, pessoas autoras e a comunidade acadêmica envolvida nas discussões sobre ciência aberta e sintetizamos os caminhos que temos trilhado nessa seara.

Palavras-chave:

Ciência aberta;
Dados abertos;
Educação em Ciências.

DATOS ABIERTOS EN LA INVESTIGACIÓN EN EDUCACIÓN EN CIENCIAS: PERSPECTIVAS, DESAFÍOS Y POSIBILIDADES

RESUMEN:

En este editorial presentamos las perspectivas de la ciencia abierta adoptadas por *Ensaio* con énfasis en prácticas de datos abiertos y datos compartidos. La disponibilidad de datos de investigaciones puede mejorar la transparencia, la colaboración, la reproducibilidad, la replicación, la reutilización y el enriquecimiento de la producción científica en el campo de la Educación en Ciencias. Por lo tanto, presentamos nuevas perspectivas acerca del tema a partir del conocimiento aprendido a lo largo de 2023 en la implantación de una sección de datos, cuyo trabajo detallamos en el flujograma. Además, abordamos algunas de las posibles implicaciones del intercambio ético de datos en el área y relatamos algunos de los resultados de las interacciones que tuvimos con nuestro equipo de edición, autores y comunidad académica involucrada en las discusiones sobre los datos abiertos y también resumimos los caminos recorridos en esta esfera.

Palabras clave:

Ciencia abierta;
Datos abiertos;
Educación en Ciencias.

OPEN DATA IN SCIENCE EDUCATION RESEARCH: PERSPECTIVES, CHALLENGES, AND POSSIBILITIES

ABSTRACT:

In this editorial, we present the open science perspectives adopted by the *Ensaio* journal, emphasizing open data practices and shared data. The availability of research data can im-

Keywords:

Open science;
Open data;
Science education.

^I Universidade de Groningen, Instituto de Educação e Comunicação Científica, Groningen, Holanda.

^{II} Universidade Federal de Ouro Preto, Ouro Preto, Brasil.

prove transparency, collaboration, reproducibility, replication, reuse, and enrichment of scientific production in Science Education. We offer new perspectives on the topic based on our experiences throughout 2023 in implementing a Data Editorship, whose work we detail in a flowchart. Furthermore, we address some of the possible implications of ethical data sharing in the field and report on some of the outcomes of the interactions between our editorial team, authors, and the academic community involved in discussions on open science by summarizing the paths we have taken in this sphere.

A tendência global em direção à gestão e ao compartilhamento de dados abertos de pesquisa, permitindo o acesso e a reutilização de conjuntos de dados originais (Albagli et al., 2014; Open Knowledge International, 2013), já faz parte das diretrizes de agências de fomento à pesquisa em âmbito nacional (e.g., FAPESP) e internacional (e.g., European Commission). Esse movimento em direção aos dados abertos, iniciado no final da década de 1990, e estimulado por organizações internacionais como a Organização para a Cooperação e Desenvolvimento Econômico, OCDE (2007), e a UNESCO (2021), insere-se em uma perspectiva mais ampla de ciência aberta (Ruusalepp, 2008). Esta visão é impulsionada por uma série de vantagens, incluindo a promoção de novas pesquisas, a transparência, a reprodutibilidade, a replicação,¹ a economia de custos, a diversificação na análise e o fortalecimento da pesquisa científica aberta (Janz 2016; Chauvette et al., 2019; Mozersky et al., 2020a; Obels et al., 2020). Adicionalmente, os dados abertos contribuem para aumentar a transparência e a responsabilização, podendo resultar em uma utilização mais eficiente e econômica de dados, especialmente em pesquisas financiadas publicamente (Mozersky et al., 2020a; Prosser et al., 2022). A abertura dos dados também tem o potencial de otimizar o valor dos dados e aliviar a carga de participantes, uma vez que incentiva a reutilização de dados previamente coletados (Mozersky et al., 2020a).

Conforme um número crescente de pessoas pesquisadoras, universidades e órgãos de financiamento buscam promover a abertura de dados, é crucial reconhecer nesse debate as dimensões éticas e as dimensões ontológicas e epistemológicas que distinguem dados mais qualitativos dos quantitativos. Isso porque nem todos os dados são igualmente adequados para a abertura (Chauvette et al., 2019). A pesquisa qualitativa, altamente presente nos estudos em Educação em Ciências, apresenta desafios ontológicos, epistemológicos e metodológicos distintos, devido à sua natureza contextual e à variação na utilidade dos dados quando retirados de seus contextos. Dessa forma, é essencial promover um diálogo transparente sobre as complexidades inerentes à pesquisa qualitativa, como a preservação da qualidade e integridade dos dados. Isso se torna crucial para assegurar que a transição para dados abertos não prejudique os indivíduos voluntários das pesquisas, nem comprometa a validade e interpretação adequada desses dados na geração de conhecimento (Prosser et al., 2022).

Estamos considerando que dados são informações fundamentais coletadas ou geradas de maneira sistemática para abordar questões ou resolver problemas de pesquisa. Esses dados são a matéria-prima para o desenvolvimento de raciocínios, discussões e cálculos inerentes à pesquisa (Jaunsen, 2018). Eles podem assumir diferentes formas, dependendo do estágio de processamento. Dados brutos, por exemplo, representam a informação diretamente obtida de uma fonte, como questionários preenchidos, gravações de entrevistas, observações etnográficas, artefatos ou documentos textuais e imagéticos. Por sua vez, dados processados referem-se a informações organizadas e codificadas, mas ainda não submetidas a uma análise aprofundada, como questionários digitados e codificados, ou dados transcritos e categorizados. Já os dados tratados representam a forma mais refinada e analisada dos dados, incluindo, por exemplo, análises de conteúdo identificando

temas e subtemas, sínteses quantitativas, análises estatísticas, modelos, algoritmos e visualizações. Essa distinção entre dados brutos, processados e tratados é crucial, pois destaca a transformação de dados ao longo do ciclo de pesquisa, desde sua coleta inicial até a produção de resultados e conclusões substanciais.

Ao explorarmos o compartilhamento de dados e materiais suplementares na pesquisa, é essencial compreender que a disponibilidade desses dados pode ocorrer em diferentes fases de processamento, apresentando distintos níveis de acesso (Foster, s.d.). A noção de dados abertos refere-se a informações acessíveis a todos, sem restrições, e que podem ser licenciadas para diversas finalidades. Por outro lado, dados compartilhados permitem um amplo acesso, mas sujeito a condições como a restrição à reutilização comercial ou a necessidade de atribuição da licença de uso. Em determinadas circunstâncias, o acesso a esses dados pode ser concedido apenas a grupos específicos, como colegas de outras instituições acadêmicas. Já os dados fechados implicam restrições de confidencialidade, especialmente quando se trata de dados sensíveis. Contudo, mesmo nesses casos, é crucial divulgar os metadados das pesquisas, e o acesso pode ser controlado por meio de repositórios específicos, assegurando que somente pessoas autorizadas tenham acesso. Portanto, quando falamos de compartilhamento de dados e materiais de pesquisa, estamos também considerando esses distintos níveis de abertura, visando assegurar transparência e segurança conforme a necessidade de cada situação.

Encerramos o editorial de 2023 — *Experiências da Revista Ensaio Pesquisa em Educação em Ciências com avaliação por pares aberta* (Mendonça et al., 2023) — com a seção “*Do OPR (open peer review) aos dados abertos: nossas perspectivas*”, na qual descrevemos as ações iniciais tomadas pelo nosso corpo editorial para implementar uma política de compartilhamento aberto de dados e materiais suplementares relacionados à pesquisa. A disponibilização de dados de pesquisa é uma realidade e acreditamos que ela pode aprimorar a transparência, colaboração, replicabilidade e o enriquecimento da produção científica no campo da educação em ciências. Por isso, neste editorial, trazemos novas perspectivas à temática com base em nossos aprendizados ao longo de 2023 na implementação de uma editoria de dados, relatando alguns dos frutos das interações entre nossa equipe editorial, pessoas autoras e a comunidade acadêmica envolvida nas discussões sobre ciência aberta e sintetizando os caminhos que temos trilhado nessa seara.

QUAIS AS POSSÍVEIS IMPLICAÇÕES POSITIVAS DO COMPARTILHAMENTO DE DADOS ABERTOS?

Considerando a recente e prolífica literatura acerca de dados abertos de pesquisa, não há motivo para não sustentar, em princípio, que a prática do compartilhamento aberto de dados e materiais suplementares na pesquisa em educação em ciências não possa ser respaldada por uma série de motivos fundamentais que fortalecem a credibilidade e contribuem para a construção de novos saberes. De forma geral, pessoas pesquisadoras de diferentes áreas do conhecimento (incluindo humanidades, mas sem dados específicos sobre a área de educação) têm adotado uma perspectiva positiva frente ao compartilhamento aberto de dados de pesquisa, reconhecendo, em sua maioria, que a prática traz benefícios para a ciência (CWTS & Elsevier, 2017).

A disponibilidade de dados abertos possibilita que pessoas pesquisadoras e o público em geral examinem, avaliem e verifiquem os resultados de estudos, reduzindo, assim, as possibilidades de má conduta científica (Corti et al., 2016) e de falhas honestas. Nesse contexto, os debates em torno do compartilhamento de dados abertos incentivam pessoas pesquisadoras e instituições a adotarem padrões éticos, particularmente para evitar a manipulação de resultados, o que contribui para a integridade da pesquisa e os debates relacionados a ela. Dessa forma, a credibilidade e a transparência são frequentemente listadas como os principais motivos para compartilhar dados de pesquisa de forma aberta, uma vez que esse nível de transparência é essencial para estabelecer a confiança na pesquisa (Uhlir & Schröder, 2007).

Além disso, os dados abertos possibilitam a replicabilidade, a reprodutibilidade e a busca pela validação em pesquisas. Ao facilitar a replicação/reprodução de estudos, pode-se contribuir para a confirmação da validade dos resultados, aumentando a confiança e a robustez de estudos científicos (Uhlir & Schröder, 2007; Wilbanks et al., 2006). Apesar disso, várias áreas do conhecimento, incluindo das humanidades e biológicas, têm relatado desafios frente à chamada crise de reprodutibilidade (Freese & Peterson, 2017; Munafò et al., 2017). Sobre esse tema, ao elaborar um relatório de estudo de consenso sobre reprodutibilidade e replicabilidade na ciência, a Academia Nacional de Ciências, Engenharia e Medicina dos Estados Unidos destacou que:

The scientific enterprise depends on the ability of the scientific community to scrutinize scientific claims and to gain confidence over time in results and inferences that have stood up to repeated testing. Important throughout this process is the sharing of data and methods and the estimation, characterization, and reporting of uncertainty. Reporting of uncertainty in scientific results is a central tenet of the scientific process, and it is incumbent on scientists to convey the appropriate degree of uncertainty to accompany original claims (NAS, 2019, p.06).

Dentre os benefícios significativos do compartilhamento de dados abertos, encontramos a promoção da colaboração e inovação. Ao disponibilizar dados, especialistas podem aproveitar diversas perspectivas e abordagens alternativas para análise, frequentemente resultando em novos conhecimentos e inovações (Bispo & Kuula-Luumi, 2017). Essa prática também contribui para uma maior eficiência de recursos, pois equipes de pesquisa podem evitar a duplicação de esforços na coleta de dados semelhantes e, em vez disso, construir sobre trabalhos já existentes (Chavette et al., 2019), o que permite uma utilização mais eficaz dos recursos financeiros e humanos disponíveis, maximizando o impacto da pesquisa (Mozersky et al., 2020a).

O compartilhamento de dados abertos também possibilita a realização de estudos meta-analíticos, nos quais múltiplos conjuntos de dados podem ser combinados para uma análise abrangente (Bispo & Kuula-Luumi, 2017). Do ponto de vista metodológico, os dados abertos podem incentivar novos tipos de pesquisa e apoiar estudos sobre métodos de coleta e medição de dados, incorporando, por exemplo, ferramentas automatizadas para organização e processamento (Uhlir & Schröder, 2007). Eles também possibilitam a criação de novos conjuntos de dados, informações e conhecimentos ao combinar dados de várias fontes (Uhlir & Schröder, 2007; Wilbanks et al., 2006).

Adicionalmente, o compartilhamento aberto de dados desempenha um papel crucial no arquivamento e preservação de ativos de pesquisa. Aqui o Plano de Gestão de Dados (PGD) desempenha importante função para a condução de pesquisas, estabelecendo o percurso estratégico para coleta, organização, armazenamento e compartilhamento de dados. O PGD especifica métodos de coleta e quem terá acesso aos dados, delineando claramente protocolos de segurança e ética que serão adotados. Ao abordar questões fundamentais, como quais tecnologias serão empregadas para acessar os dados e as estruturas para armazenamento e conservação, o PGD também oferece diretrizes detalhadas para garantir a integridade e segurança dos dados ao longo do tempo. Embora a elaboração do PGD ainda não seja uma prática corrente em muitas áreas de pesquisa (Gorgolewski et al., 2013), sua relevância acompanha as discussões sobre dados abertos e sua elaboração têm crescido entre agências de fomento nacionais e internacionais.²

De forma ilustrativa, trazemos como exemplo um estudo recente que destacou a acentuada queda na acessibilidade dos dados de 516 estudos ecológicos publicados entre 1991 e 2011 (Vines et al., 2014). O estudo relatou que a probabilidade de acesso a esses dados diminuiu 17% ao ano, com apenas 20% dos estudos da década de 1990 mantendo seus dados disponíveis. Encontrar pessoas autoras e receber respostas foi desafiador, com sucesso em apenas 37% dos casos, e a probabilidade de encontrar endereços de e-mail funcionais diminuiu 7% ao ano (Vines et al., 2014). Levantamentos mais recentes em outros contextos, como na ciência política, também registraram dificuldades de acesso a dados e materiais complementares de estudos publicados bem como na replicação desses estudos (Avelino et al., 2021).

O compartilhamento aberto de dados serve não apenas como um “cofre” para o conhecimento científico, mas também oferece a oportunidade de aplicar novos métodos de análise e explorar tópicos que podem não ter sido previstos anteriormente (Uhlir & Schröder, 2007), o que incentiva novas pesquisas e novos tipos de pesquisa, enriquecendo não apenas a base de conhecimentos, mas também abrindo caminhos para inesperadas maneiras de compreender fenômenos e objetos de estudo. Além disso, a disponibilidade de dados abertos está em sintonia com a filosofia da educação aberta, que visa tornar o conhecimento acessível. Isso é de particular importância para a formação de novas equipes de pesquisa, que passam a ter acesso a dados reais para aprender e praticar abordagens, estratégias e ferramentas metodológicas, enriquecendo a aprendizagem. Corroborando essa perspectiva, um estudo sobre o uso de repositórios de dados no Reino Unido e na Finlândia destacou que estudantes e funcionários de instituições de ensino são os maiores usuários de repositórios de dados abertos de natureza qualitativa (Bispo & Kuula-Luumi, 2017).

Numa perspectiva otimista e em termos teóricos, considera-se também que dados abertos de pesquisa podem contribuir para a transferência de informações, promovendo o desenvolvimento de capacidades em países com menos recursos para pesquisa e incentivando a pesquisa interdisciplinar, intersectorial, interinstitucional e internacional (Uhlir & Schröder, 2007; Wilbanks et al., 2006). Por fim, assume-se por princípio que a prática do compartilhamento de dados como parte da ciência aberta amplia o retorno do investimento público em pesquisa para a sociedade.

ESPECIFICIDADES DA PESQUISA QUALITATIVA E CONSIDERAÇÕES SOBRE O (NÃO) COMPARTILHAMENTO DE DADOS

Apesar dos pontos que sintetizamos na seção anterior, ao debater sobre o compartilhamento de dados é fundamental reconhecer explicitamente as peculiaridades inerentes à pesquisa tradicionalmente qualitativa em Educação em Ciências, bem como as considerações que podem restringir a viabilidade do compartilhamento desses dados. A literatura sobre a temática em Educação é ainda escassa,³ o que nos instiga a propor um diálogo interdisciplinar com pessoas pesquisadoras provenientes de outras áreas e países que têm se debruçado sobre o tema há mais tempo. Merece destaque o fato de que as políticas e recursos relativos ao compartilhamento de dados abertos exibem variações consideráveis em âmbito global, porém, a maioria delas parece ter sido desenvolvida com o propósito de atender a dados de natureza quantitativa (Mozersky et al., 2020a).

Não obstante, é parte intrínseca das discussões acerca da abertura de dados o princípio de que estes devem ser acessíveis sempre que possível, porém, devem permanecer restritos quando necessário, preservando, contudo, a abertura dos metadados, que servem para contextualizar a natureza dos dados. Assim, o notável crescimento do movimento de ciência aberta sugere que os estudos qualitativos também podem estar sujeitos a requisitos crescentes no que tange ao compartilhamento de dados (Bishop & Kuula-Luumi, 2017; Crosas et al., 2018). Cumpre notar que isso não implica em um compartilhamento indiscriminado e antiético dos dados, pois há distintos níveis de abertura de dados sob a perspectiva da ciência aberta como mencionamos anteriormente, e os dados podem ser disponibilizados em distintas fases do processo de pesquisa (Branney et al., 2019). A seguir, apresentamos brevemente alguns aspectos que permeiam esse debate, visando ampliar o diálogo, dado que esta é uma discussão em curso, ainda incipiente na pesquisa em Educação em Ciências.

No contexto europeu e australiano, observamos a existência de políticas consolidadas para estimular o compartilhamento de dados qualitativos, ao passo que nos Estados Unidos e no Brasil, a prática é relativamente recente. Dentre as razões que explicam a resistência ao compartilhamento de dados qualitativos incluem-se a natureza contextual e sensível desses dados, a carência de apoio institucional e reconhecimento, financiamento adequado e a ausência de infraestrutura e tecnologias necessárias para a gestão de dados (Tenopir et al., 2011; CWTS & Elsevier, 2017; Mozersky et al., 2020a). Desafios relacionados ao tempo e à atua-

lização contínua também representam obstáculos substanciais. Assim, aqui entendemos que qualquer paradigma de compartilhamento de dados leve em consideração tanto as dificuldades enfrentadas pelas pessoas pesquisadoras na geração de dados quanto o valor agregado para usuários finais (Jackson & Pachter, 2023).

As dimensões de apoio e reconhecimento às pessoas pesquisadoras que adotam como prática o compartilhamento de dados têm sido também uma das pautas das discussões sobre ciência aberta, haja vista que contribuições e cargas de trabalho podem ser reconhecidas de forma desigual entre diferentes áreas do conhecimento. Nesse contexto, a atribuição de um Digital Object Identifier (DOI) aos conjuntos de dados depositados no Dataverse pode ser considerada um estímulo significativo para a comunidade científica, visto que proporciona identificação única e permanente, facilitando citações precisas, rastreamento do impacto ao longo do tempo e promovendo a reutilização eficiente dos dados. Apesar disso, no contexto brasileiro ainda não há campos específicos designados para esse propósito na plataforma nacional Lattes, ainda que a inclusão desses conjuntos de dados seja viável no campo de Produção Técnica.

Do ponto de vista ético, é crucial destacar que os dados qualitativos frequentemente abordam questões sensíveis e estigmatizadas que raramente podem ser exploradas de maneira abrangente por meio de abordagens quantitativas. Isso gera consideráveis preocupações relativas ao consentimento informado, à propriedade dos dados, à confidencialidade e à necessária preservação do anonimato (DuBois et al., 2018; Prosser et al., 2022). A possibilidade de reidentificação de participantes, sobretudo em virtude do detalhamento intrínseco dos dados qualitativos, pode acarretar implicações adversas, tornando o compartilhamento de tais dados um terreno ético delicado e complexo (Mozersky et al., 2020b).

Visando aprofundar a questão, Mozersky et al. (2020b) conduziram um estudo nos Estados Unidos no qual realizaram entrevistas com 30 indivíduos que haviam participado de estudos qualitativos sensíveis, a fim de explorar suas perspectivas e preocupações quanto ao compartilhamento de dados qualitativos. As pessoas participantes se mostraram favoráveis ao compartilhamento dos dados, desde que houvesse procedimentos de desidentificação e que o compartilhamento fosse conduzido por especialistas. Os indivíduos levantaram inquietações relativas à confidencialidade e ao uso inadequado dos dados, mas, em sua maioria, expressaram a expectativa de que seus dados fossem compartilhados, demonstrando uma possível maior obrigação de compartilhamento do que anteriormente se considerava (Mozersky et al., 2020b). É importante notar que estudos desse teor ainda são raros, sobretudo no que concerne a populações vulneráveis, incluindo crianças, que frequentemente figuram como participantes em pesquisas no campo da Educação em Ciências. Assim, à medida que a discussão sobre a ética de dados qualitativos avança, a promoção de práticas mais éticas no compartilhamento e uso desses dados é fundamental para equilibrar a necessidade de transparência e a proteção de participantes, especialmente em contextos de pesquisa sensível.

A natureza contextual dos dados qualitativos apresenta um desafio epistemológico significativo. As pessoas pesquisadoras que originalmente coletaram esses dados detêm um profundo entendimento do contexto e da interpretação subjacentes. Portanto, a análise secundária de dados qualitativos realizada por pessoas pesquisadoras que não possuem a expertise contextual e a experiência de especialistas que originaram os dados originais pode ser uma tarefa complexa. Não obstante, nem todos os dados qualitativos são inviáveis para compartilhamento. Recursos como documentos de metadados detalhados, livros de códigos, guias de entrevista e transcrições anonimizadas podem ser disponibilizados para permitir a verificação das conclusões alcançadas em estudos qualitativos específicos. Além disso, análises secundárias desses dados podem resultar na obtenção de novas percepções, enriquecendo o campo da pesquisa qualitativa (DuBois et al., 2018) e têm também se tornado uma prática crescente, inclusive com a criação de repositórios específicos para depósito de dados qualitativos,⁴ como abordamos na seção a seguir.

DADOS COMPARTILHADOS SÃO REUTILIZADOS?

Há diversas formas de compartilhar dados de pesquisa, e o Dataverse desponta como uma dentre as plataformas apropriadas para esse propósito. Amplamente utilizado em instituições acadêmicas e científicas, o Dataverse desempenha um papel fundamental na promoção de uma cultura de dados abertos, facilitando o compartilhamento, a publicação e a gestão de conjuntos de dados científicos. Tal plataforma, de código aberto, consente às instituições adotá-la, personalizá-la e contribuir com seu desenvolvimento, ofertando uma infraestrutura centralizada destinada a armazenar, ordenar, e disponibilizar dados. Ademais, ela provê informações detalhadas, tais como metadados e documentação, tornando os dados acessíveis tanto à comunidade acadêmica quanto ao público em geral (King, 2007). Adicionalmente, é notável o uso do Dataverse por entidades como a Scientific Electronic Library Online (SciELO), destacando sua importância no âmbito do ecossistema da pesquisa científica e na disseminação de conhecimento no contexto do Sul Global.

Dados apresentados pelo projeto Dataverse⁵ em dezembro de 2023 indicavam que existiam 15,9 mil Dataverses, abrangendo 403 mil conjuntos de dados (datasets) e quase 5 milhões de arquivos digitais. Dois conjuntos de informações derivados desses dados se destacam e suscitam reflexões para o diálogo que estamos propondo neste editorial. Em primeiro lugar, o número total de downloads desses arquivos, os quais podem servir como um indicador inicial, ainda que genérico, do interesse despertado por estes dados. No mês de dezembro de 2023, houve cerca de 74,8 milhões de downloads, considerando o conjunto de arquivos depositados nos Dataverses monitorados. Em segundo lugar, é pertinente salientar o compartilhamento de dados por áreas do conhecimento. Das coleções de datasets, as áreas de Ciências Sociais e Artes e Humanidades, tradicionalmente associadas à pesquisa qualitativa, foram responsáveis por cerca de 62% dos datasets compartilhados nos Dataverses monitorados pelo projeto.⁶

Até o momento, carecemos de informações detalhadas acerca da utilização desses repositórios de dados para fins de pesquisa, especialmente na área da Educação. Trabalhos avaliando citações de datasets compartilhados abertamente poderiam ser um bom indicativo sobre o uso (e a qualidade) desses dados. Dessa forma, assim como fizemos nas seções precedentes deste editorial, trazemos reflexões e estudos provenientes de outras áreas do conhecimento como forma de estimular e aprofundar a discussão acerca do compartilhamento e utilização de dados abertos no âmbito da pesquisa em Educação em Ciências.

Como anteriormente destacado, as implicações decorrentes da adoção de dados abertos em pesquisas qualitativas envolvem questões de cunho epistemológico, metodológico, ético (Alexander et al., 2019; Childs et al., 2014) e também jurídico (ver Prosser et al., 2022). Mais especificamente, surgem desafios quanto à análise secundária, que implica a reutilização de dados qualitativos previamente coletados em estudos anteriores. Tal prática, como sumarizamos, pode ser motivada pela busca de transparência, validação de resultados e pela criação de novos saberes a partir de dados preexistentes.

Um dos principais desafios associados ao uso de dados abertos em pesquisas qualitativas reside na reutilização fora do contexto original. Os dados qualitativos encapsulam experiências individuais por meio de palavras, imagens ou comportamentos, sendo moldados por observações, entrevistas, documentos e artefatos (Chauvette et al., 2019; Weller, 2023). A singularidade dos dados qualitativos emerge de contextos históricos, culturais e sociais, conferindo-lhes uma natureza inimitável (Creswell & Poth, 2018; Weller, 2023). A relação entre pessoas pesquisadoras e participantes conduz à construção conjunta do conhecimento, realçando, assim, a importância do contexto dos dados. Adicionalmente, dados qualitativos são intrinsecamente subjetivos, influenciados pelas experiências e perspectivas tanto das pessoas participantes quanto das próprias pessoas pesquisadoras. Dessa forma, a reutilização de dados qualitativos sem a devida contextualização pode prejudicar a interpretação e o significado destes dados.

A natureza dos métodos específicos empregados em pesquisas qualitativas também traz desafios à reutilização de dados abertos. Alguns delineamentos de pesquisa não se prestam à análise secundária, devido à forte influência de especialistas e à natureza contextual das experiências de pessoas participantes. Em pesquisas participativas, por exemplo, os dados não se limitam a simples transcrições, o que dificulta a reutilização sem o envolvimento de pessoas participantes (Chauvette et al., 2019).

As complexidades éticas e legais relacionadas ao compartilhamento de dados abertos em pesquisas qualitativas não podem ser subestimadas. O consentimento informado de participantes, bem como as questões de privacidade e confidencialidade figuram como elementos cruciais (Jesus, 2019). O compartilhamento de dados pode entrar em conflito com as condições originais sob as quais o consentimento foi obtido. Ademais, a identificação inadvertida de participantes devido a detalhes específicos de suas experiências é motivo de preocupação. Em comunidades pequenas ou contextos específicos, a identidade de participantes pode ser reconstituída, comprometendo a privacidade (Chauvette et al., 2019). A proteção da identidade de indivíduos e o anonimato podem ser desafiantes, podendo resultar na exclusão de informações cruciais, tornando os dados inutilizáveis.

Diante destas ponderações substanciais, surge a questão sobre a pertinência do compartilhamento de dados qualitativos para reuso em contextos diferentes daqueles nos quais foram originalmente concebidos. Apesar das aparentes limitações, repositórios dedicados à disponibilização de dados qualitativos têm experimentado crescimento, sendo uma realidade em países como Alemanha, Austrália, Eslovênia, Finlândia, Irlanda, Noruega, Polônia, e Reino Unido (Bishop & Kuula-Luumi, 2017).

Ao aprofundar essa temática, Bishop e Kuula-Luumi (2017) conduziram um estudo focando em dois repositórios no Reino Unido e na Finlândia, constatando um aumento no uso de dados qualitativos para análise secundária. O estudo demonstrou que a maioria das coleções de dados foi utilizada, refutando a premissa de que os dados eram apenas arquivados nos repositórios sem jamais serem empregados. Além disso, a pesquisa apontou que (i) os dados primários podem se destacar como recursos valiosos para abordar questões complexas, quando as respostas diretas são de difícil obtenção devido à delicadeza dos tópicos e (ii) uma variedade de abordagens para análise qualitativa secundária emergiu e se expandiu com o tempo.

Além do estudo de Bishop e Kuula-Luumi (2017), há evidências da expansão da reutilização de dados em análise qualitativa secundária em disciplinas inteiras, como tem sido debatido de forma ampla na área da Psicologia nos últimos anos (DuBois et al., 2018). Para algumas pessoas autoras, o compartilhamento e o uso de dados qualitativos já se tornaram uma realidade, ainda que pessoas pesquisadoras permaneçam inconscientes ou resistentes a tais discussões (Prosser et al., 2022; DuBois et al., 2018; Strasser, 2013). Assim, entendemos que a continuidade desses diálogos e a promoção de melhores práticas no compartilhamento e uso de dados qualitativos são essenciais para avançar na compreensão das implicações e benefícios desse processo em diversas áreas da pesquisa acadêmica, incluindo a Educação em Ciências.

CONSIDERAÇÕES SOBRE O COMPARTILHAMENTO DE DADOS NA PESQUISA EM EDUCAÇÃO EM CIÊNCIAS COM BASE NOS PRINCÍPIOS FAIR E CARE

O compartilhamento e reuso de dados na pesquisa científica dependem, em grande parte, de dois fatores essenciais: a adequada arquivagem dos dados e a capacidade de encontrá-los posteriormente. Nesse contexto, os metadados e os arquivos “readme” funcionam como chaves para desbloquear o potencial dos dados. Eles operam como guias que descrevem o conteúdo, a estrutura e o contexto dos conjuntos de dados, tornando-os compreensíveis para outras pessoas pesquisadoras e partes interessadas. Quando os metadados e

arquivos “readme”⁷ estão bem elaborados, possibilitam a descoberta e utilização eficaz dos dados, garantindo que a informação seja clara, estruturada e mais prontamente acessível (Elger et al., 2016; SciELO, 2021, 2023).

Neste ponto, entram em cena os princípios FAIR, que estabelecem um conjunto de diretrizes que transcendem os arquivos “readme” tradicionais. Eles aumentam a garantia de que os dados sejam encontrados facilmente (Findable), acessíveis a pessoas interessadas (Accessible), interoperáveis entre diferentes sistemas e ferramentas (Interoperable) e reutilizáveis em várias pesquisas (Reusable) (Wilkinson et al., 2016). Esses princípios têm sido amplamente adotados para melhorar a eficiência na utilização de dados científicos no arquivamento em repositórios como os Dataverses.

Os princípios FAIR representam um avanço qualitativo no compartilhamento de dados, pois não apenas tornam os dados “legíveis”, mas também facilitam a colaboração e a solução de desafios complexos em diversas disciplinas. Ao seguir os princípios FAIR, pessoas pesquisadoras e instituições promovem a transparência e a colaboração, alinhando-se com práticas estabelecidas de gestão de dados na era digital (Wilkinson et al., 2016). No entanto, os princípios FAIR, embora essenciais, não abordam completamente as complexidades inerentes aos dados de pesquisa em áreas sensíveis, como a Educação em Ciências. Questões éticas e culturais desempenham um papel fundamental na coleta e no uso desses dados, e é importante que princípios associados ao cuidado e respeito com esses dados também sejam considerados.

Nesse contexto, iniciativas como o “The Good Data Manifesto” (Trenham & Steer, 2019) e os princípios CARE (Collective Benefit, Authority to Control, Responsibility, Ethics) (GIDA, 2018) têm ganhado destaque. Eles buscam garantir o uso responsável de dados, considerando o impacto social, cultural e ético (Borsetti et al., 2021). Os princípios CARE, em particular, focam na autoridade das comunidades sobre seus dados, promovendo a ética e a responsabilidade no tratamento e uso dessas informações (GIDA, 2018).

Essa combinação de princípios FAIR e CARE tem formado o movimento para o compartilhamento de dados orientado pelo mote #SejaJUSTOeCUIDADOSO (#BeCARE&FAIR), que visa promover uma abordagem inclusiva e ética para o compartilhamento de dados, respeitando as diversas necessidades e contextos das populações envolvidas, garantindo que os dados sejam usados de forma respeitosa e benéfica, levando em consideração suas línguas, culturas e visões de mundo (Borsetti et al., 2021). Isso assegura que o compartilhamento de dados na pesquisa em campos sensíveis como a educação não apenas atenda a critérios técnicos, mas também respeite princípios éticos e culturais essenciais. Nesse contexto, os arquivos “readme” são um ponto de partida, os princípios CARE representam os alicerces éticos e responsáveis necessários para lidar com informações sensíveis, e os princípios FAIR são o caminho pragmático que possibilita o compartilhamento e reuso de dados na pesquisa.

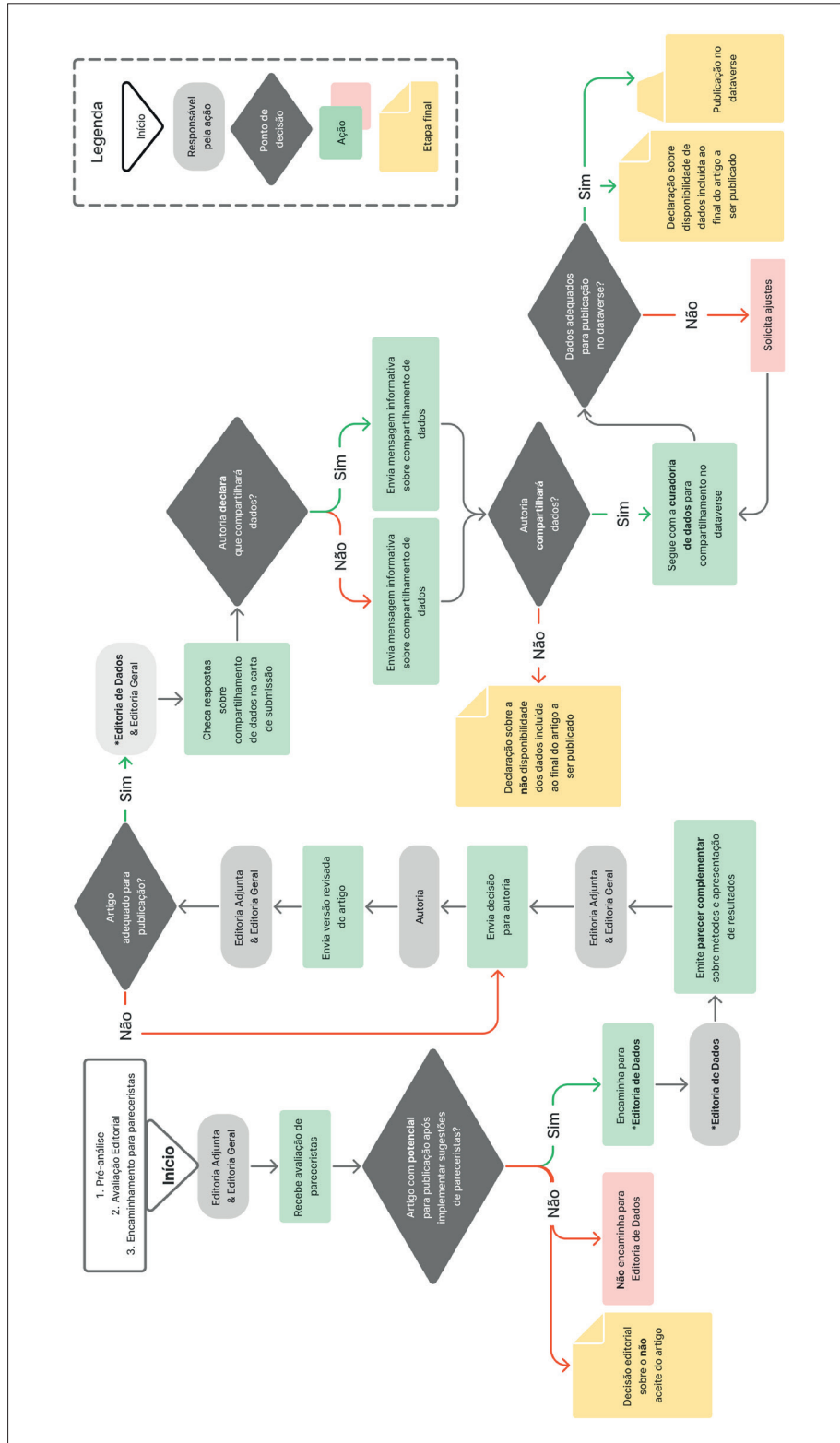
CAMINHOS TRILHADOS PELA ENSAIO: PESQUISA EM EDUCAÇÃO E CIÊNCIAS RUMO AO COMPARTILHAMENTO DE DADOS E MATERIAIS SUPLEMENTARES À PESQUISA

Neste editorial, apresentamos uma ampliação dos temas anteriormente abordados em nossos editoriais sobre práticas de Ciência Aberta (ver Mendonça & Franco, 2021; Mendonça et al., 2023). Essa expansão foi promovida por meio de reuniões formativas periódicas entre o corpo editorial da revista, durante as quais realizamos leituras e discussões aprofundadas sobre questões cruciais, como diretrizes para a Ciência Aberta e ética na pesquisa em educação, propostas em materiais organizados pela Associação Nacional de Pós-Graduação e Pesquisa em Educação (ANPED). Aprofundar nossa compreensão desses tópicos tem sido crucial para direcionar nossa abordagem em prol de uma Ciência Aberta mais responsável e ética nas ações editoriais.

Outra realização notável foi a expansão dos conjuntos de dados publicados no Dataverse da Ensaio: Pesquisa em Educação e Ciências (EPEC), organizados de acordo com as diretrizes da SciELO (SciELO, 2021, 2023). Essa colaboração contínua com a equipe do SciELO Data tem possibilitado um diálogo construtivo e valioso, contribuindo para o avanço do compartilhamento de dados em nossa área de pesquisa e auxiliado nosso corpo editorial a melhor apoiar as pessoas autoras que buscam publicar na Ensaio.

Adicionalmente, estabelecemos e estruturamos uma Editoria de Dados dedicada a revisar e aprimorar constantemente nossa política editorial para compartilhamento de dados (Figura 1). Suas atividades englobam desde a avaliação da possibilidade de compartilhamento de dados, incluindo a definição do nível de abertura e a forma de disponibilização, até a verificação de eventuais pendências na comunicação da parte metodológica e de representação dos dados. Além disso, a Editoria de Dados emite eventuais pareceres complementares à Editoria Adjunta e, quando oportuno, convida as pessoas autoras a compartilharem seus dados. Posteriormente, realiza a curadoria dos dados e presta auxílio às pessoas autoras no processo de compartilhamento, culminando na publicação dos dados no Dataverse da revista.⁸

Figura 1. Fluxograma do processo editorial simplificado da EPEC, destacando as ações da Editoria de Dados



Por meio dessa editoria, temos buscado manter uma comunicação próxima e constante com as pessoas autoras ao longo de todo o processo de editoração, buscando compreender suas dúvidas e inquietações. Essa interação contínua não apenas enriqueceu nossa compreensão das complexidades da Ciência Aberta, mas também serviu como uma fonte de inspiração para a elaboração deste editorial. Nesse processo, percebemos um grande potencial para aprofundar nossa compreensão sobre esse tema e buscamos contribuir para fortalecer as conversas sobre essa temática para além das páginas deste periódico. Em adição a este editorial, elaboramos recursos adicionais, como “Guia para preparação de conjunto de dados para depósito em repositório: orientações para a área de Educação em Ciências”, para orientar a preparação de conjuntos de dados, facilitando sua submissão ao Dataverse da revista.

Nosso compromisso com a Ciência Aberta transcende as fronteiras da revista. Participamos ativamente em eventos promovidos pela SciELO, como convidados para debates de relevância, a exemplo do Seminário SciELO 25 Anos, onde discutimos políticas editoriais de dados em periódicos da SciELO Brasil.⁹ Além disso, também investimos na criação de módulos dedicados à Ciência Aberta e à ética na pesquisa em educação, integrados em disciplinas oferecidas a estudantes de pós-graduação em universidades públicas (como UFOP e UFMG). Essa iniciativa visa preparar uma nova geração de pessoas pesquisadoras e editoras científicas equipadas com as ferramentas e os princípios necessários para promover práticas de pesquisa e publicação mais abertas e éticas.

Diante das reflexões apresentadas neste editorial e da escassez de discussões na área sobre dados abertos na pesquisa em Educação em Ciências, reiteramos a importância de promover um debate aberto e abrangente sobre esse tópico, em diversos espaços e formatos. Convidamos a comunidade de Educação em Ciências a se juntar a esse diálogo. As implicações derivadas das pesquisas apresentadas aqui enfatizam a urgência de adotar políticas de compartilhamento de dados abertos, não apenas como uma salvaguarda para informações valiosas, mas também como um catalisador para o desenvolvimento sustentável do conhecimento em nossa área. Complementarmente, para sustentar essa crescente tendência em direção aos dados abertos qualitativos, é fundamental que as implicações epistemológicas, metodológicas, legais e éticas sejam abordadas de maneira apropriada, fornecendo uma base sólida para essa prática emergente.

Inspirados pelas discussões já em andamento em áreas como a Psicologia (por exemplo, DuBois et al., 2018; Roller & Lavrakas, 2018), consideramos que o impacto significativo do compartilhamento de dados na área de Educação em Ciências dependerá, em grande parte, das decisões tomadas por pessoas pesquisadoras quanto aos materiais e recursos que escolherem compartilhar de maneira aberta, o que pode resultar em benefícios metodológicos notáveis em comparação com um cenário no qual o compartilhamento de dados é escasso. Para finalizar, esperamos que o Editorial e o Material Complementar tenham contribuído para esclarecer sobre tipos, características e finalidades de dados abertos na pesquisa e encorajar a prática ética e responsável de compartilhamento de dados.

AGRADECIMENTOS

A Universidade Federal de Ouro Preto, pelos recursos para tradução do editorial nas línguas inglesa e espanhola. Ao CNPQ pela bolsa de produtividade em pesquisa da 2ª autora.

DECLARAÇÃO DE AUTORIA

NATHÁLIA HELENA AZEVEDO: Análise formal, Conceituação, Redação – Primeira versão, Redação – Revisão e Edição, Visualização.

PAULA CRISTINA CARDOSO MENDONÇA: Análise formal, Redação – Revisão e Edição, Financiamento, Visualização.

REFERÊNCIAS

- Albagli, S., Clinio, A., & Raychtock, S. (2014). Ciência Aberta: correntes interpretativas e tipos de ação. *Liinc em revista*, 10(2).
- Alexander, S. M., Jones, K., Bennett, N. J., Budden, A., Cox, M., Crosas, M., ... Weber, N. (2019). Qualitative data sharing and synthesis for sustainability science. *Nature Sustainability*. doi:10.1038/s41893-019-0434-8
- Avelino, G., Desposato, S., & Mardegan, I. (2021). Transparency and Replication in Brazilian Political Science: A First Look. *Dados*, 64.
- Bishop, L., & Kuula-Luomi, A. (2017). Revisiting qualitative data reuse: A decade on. *SAGE open*, 7(1), 2158244016685136.
- Borsetti, S., Vidotti, G., Torino, E., & Coneglian, C. (2021). #SejaJUSTOeCUIDADOSO: princípios FAIR e CARE na gestão de dados de pesquisa. In: SALES, Luana Farias; VEIGA, Viviane dos Santos; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). *Princípios FAIR aplicados à gestão de dados de pesquisa*. Rio de Janeiro: Ibict, 2021. p. 201-214.
- Branney, P., Reid, K., Frost, N., Coan, S., Mathieson, A., & Woolhouse, M. (2019). A context-consent meta-framework for designing open (qualitative) data studies. *Qualitative Research in Psychology*, 16(3), 483-502.
- Chauvette, A., Schick-Makaroff, K., & Molzahn, A. E. (2019). Open data in qualitative research. *International Journal of Qualitative Methods*, 18, 1609406918823863.
- Childs S., McLeod J., Lomas E., Cook G. (2014). Opening research data: Issues and opportunities. *Records Management Journal*, 24, 142.
- Corti L., Fielding N., Bishop L. (2016). Editorial for special edition, digital representations: Re-using and publishing digital qualitative data. *SAGE Open*, 6, 1.
- Creswell J. W., Poth C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Crosas, M., Gautier, J., Karcher, S., Kirilova, D., Otalora, G., & Schwartz, A. (2018). Data policies of highly-ranked social science journals. Retrieved from https://dataverse.org/sites/projects.iq.harvard.edu/files/dataverseorg/files/data_policies_of_highly-ranked_social_science_journals.pdf
- Centre for Science and Technology Studies (CWTS) & Elsevier. (2017). Open Data: the researcher perspective. https://www.elsevier.com/___data/assets/pdf_file/0004/281920/Open-data-report.pdf
- DuBois, J. M., Strait, M., & Walsh, H. (2018). Is it time to share qualitative research data? *Qualitative Psychology*, 5(3), 380–393.
- Elger, K., Biskaborn, B. K., Pampel, H., & Lantuit, H. (2016). Open research data, data portals and data publication—an introduction to the data curation landscape. *Polarforschung*, 85(2), 119-133.
- Foster. (s.d.). *Managing and Sharing Research Data*. <https://www.fosteropenscience.eu/>
- Freese, J., & Peterson, D. (2017). Replication in social science. *Annual Review of Sociology*, 43, 147-165.
- GIDA, The Global Indigenous Data Alliance. (2018). CARE principles for Indigenous data governance <http://gida-global.org/>

- Gorgolewski, K. J., Margulies, D. S., & Milham, M. P. (2013). Making data sharing count: a publication-based solution. *Frontiers in neuroscience*, 7, 9.
- Jackson, K. C., & Pachter, L. (2023). A standard for sharing spatial transcriptomics data. *Cell Genomics*, 3(8).
- Janz, N. (2016). Bringing the gold standard into the classroom: replication in university teaching. *International Studies Perspectives*, 17(4), 392-407.
- Jaunsen, A. (2018). The state of Open Science in the Nordic countries: Enabling data-driven science in the Nordic countries. Nordic Co-operation. <http://norden.diva-portal.org/smash/record.jsf?pid=diva2%3A1257306&dsid=3471>
- Jesus, D. M. (2019). Integridade na coleta, na produção e na análise de dados. In ANPED. *Ética e pesquisa em Educação: subsídios*, vol 1. 127-132.
- King, G. (1995). Replication, Replication. *PS: Political Science and Politics*, vol. 28, n.03, pp. 444-452.
- King, G. (2007). An Introduction to the Dataverse Network as an Infrastructure for Data Sharing. *Sociological Methods and Research*, 36, Pp. 173–199.
- Mendonça, P. C. C., Franco, L. G. (2021). Editorial: A ciência aberta e a área de Educação em Ciências – perspectivas e diálogos. *Ensaio Pesquisa em Educação em Ciências*, 23, <https://doi.org/10.1590/1983-21172021230102>
- Mendonça, P. C. C., Franco, L. G., Massi, L., & Coelho, G. R. (2023). Editorial: Experiências da revista Ensaio Pesquisa em Educação em Ciências com avaliação por pares aberta. *Ensaio Pesquisa em Educação em Ciências*, 25, <https://doi.org/10.1590/1983-21172022240137>
- Mozerky, J., Walsh, H., Parsons, M., McIntosh, T., Baldwin, K., & DuBois, J. M. (2020a). Are we ready to share qualitative research data? Knowledge and preparedness among qualitative researchers, IRB members, and data repository curators. *IASSIST quarterly*, 43(4).
- Mozerky, J., Parsons, M., Walsh, H., Baldwin, K., McIntosh, T., & DuBois, J. M. (2020b). Research participant views regarding qualitative data sharing. *Ethics & human research*, 42(2), 13-27.
- Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., ... & Ioannidis, J. (2017). A manifesto for reproducible science. *Nature human behaviour*, 1(1), 1-9.
- National Academies of Sciences, Engineering, and Medicine. (2019). Reproducibility and replicability in science.
- Organization for Economic Cooperation and Development. (2007). OECD principles and guidelines for access to research data from public funding. Retrieved from <http://www.oecd.org/sti/sci-tech/38500813.pdf>
- Open Data Handbook. (2013). What is open data? *Open Knowledge International*. Retrieved from <http://opendatahandbook.org/guide/en/what-is-open-data/>
- Obels, P., Lakens, D., Coles, N. A., Gottfried, J., & Green, S. A. (2020). Analysis of open data and computational reproducibility in registered reports in psychology. *Advances in Methods and Practices in Psychological Science*, 3(2), 229–237.
- Prosser, A.M., Hamshaw, R.J., Meyer, J., Bagnall, R., Blackwood, L., Huysamen, M., ... & Walter, Z. (2022). When open data closes the door: A critical examination of the past, present and the potential future for open data guidelines in journals. *British Journal of Social Psychology*.
- Roller, M. R., & Lavrakas, P. J. (2018). A total quality framework approach to sharing qualitative research data: Comment on Dubois et al. (2018). *Qualitative Psychology*, 5(3), 394–401.
- Ruusalepp R. (2008). A comparative study of international approaches to enabling the sharing of research data. Retrieved from https://www.era.lib.ed.ac.uk/bitstream/handle/1842/3361/Ruusalepp%20Data_Sharing_Report.pdf;jsessionid=AEA3F674EFD77EB4667BA978EC45EE48?sequence=1
- SciELO. Guia de preparação de dados de pesquisa [online]. SciELO, 2021. Available from https://wp.scielo.org/wp-content/uploads/Guia_preparacao_pt.pdf

SciELO. Guia de depósito de dados de pesquisa [online]. SciELO, 2023. Available from https://wp.scielo.org/wp-content/uploads/Guia_deposito_pt.pdf

Strasser, C. (2013). Some answers to the most common misconceptions about sharing research data. London School of Economics. <https://blogs.lse.ac.uk/impactofsocialsciences/2013/05/06/answers-to-common-misconceptions-about-data-sharing/>

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E., Manoff, M. & Frame, M. (2011). Data sharing by scientists: practices and perceptions, PLoS One, 6, e21101.

Trenham, C., & Steer, A. (2019). The good data manifesto. In Angela Daly, S. Kate Devitt and Monique Mann. Good Data, 37-53.

Uhler, P.F., & Schröder, P. (2007). Open Data for Global Science. Data Science

UNESCO (2021). UNESCO Recommendation on Open Science. <https://doi.org/10.54677/MNMH8546>

van Dijk, W., Schatschneider, C., & Hart, S. A. (2021). Open science in education sciences. Journal of Learning Disabilities, 54(2), 139-152.

Vines, T. H., Albert, A. Y., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., ... & Rennison, D. J. (2014). The availability of research data declines rapidly with article age. Current biology, 24(1), 94-97.

Weller, S. (2023). Fostering habits of care: Reframing qualitative data sharing policies and practices. Qualitative Research, 23(4), 1022-1041.

Wilbanks, J., Boyle, J., Reynolds, W.N. (2006). Introduction to Science Commons. Retrieved from http://sciencecommons.org/wpcontent/uploads/ScienceCommons_Concept_Paper.pdf

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, 3(1), 1-9.

NOTAS

1 Os termos transparência, reprodutibilidade e replicação frequentemente surgem como intercambiáveis na literatura. Neste texto, seguimos a definição proposta por Janz (2016) e King (1995), em que a transparência se refere à disponibilização dos materiais de pesquisa para a comunidade acadêmica, a reprodutibilidade diz respeito à capacidade de reproduzir um estudo utilizando os dados disponíveis, enquanto a replicação é entendida como a capacidade de reproduzir um estudo considerando novas variáveis ou hipóteses.

2 Há uma série de recursos e ferramentas online para facilitar a criação de um PGD. A FAPESP já realizou uma curadoria da temática, disponível em <https://fapesp.br/gestaodedados-documentosinteress>

3 Aqui, destacamos trabalhos relacionados à ciência da aprendizagem, como o de van-Dijk et al. 2021

4 Por exemplo (i) <https://discover.ukdataservice.ac.uk/qualibank> e (ii) <https://qdr.syr.edu/>.

5 <https://dataverse.org/>

6 Projetos internacionais que monitoram outros repositórios na Europa e Estados Unidos (e.g., <https://www.re3data.org/>) indicam a existência de repositórios específicos dedicados às Ciências da Vida e da Natureza, bem como repositórios que indicam aceitar dados da área de Educação.

7 Complementando este editorial, elaboramos o Guia para preparação de conjunto de dados para depósito em repositório: orientações para a área de Educação em Ciências, no qual abordamos em detalhes as características técnicas dos datasets e do documento readme. O guia está disponível no dataverse da revista <https://data.scielo.org/dataverse/brepec>

8 <https://data.scielo.org/dataverse/brepec>

9 <https://25.scielo.org/seminarios/scielo-data/>

Nathália Helena Azevedo

Pesquisadora de pós-doutorado (Instituto de Educação e Comunicação Científica), Universidade de Groningen, Groningen, Holanda. Doutora em Educação em Ciências (Programa Interunidades em Educação em Ciências, USP, Brasil).

E-mail: helena.nathalia@gmail.com

Paula Cristina Cardoso Mendonça

Professora Associada na Universidade Federal de Ouro Preto, Ouro Preto, Brasil. Doutora em Educação (Universidade Federal de Minas Gerais, UFMG, Brasil).

E-mail: paulamendonca@ufop.edu.br

Contato:

Centro de Ensino de Ciências e Matemática de Minas Gerais – CECIMIG

Faculdade de Educação | Universidade Federal de Minas Gerais

Av. Pres. Antônio Carlos, 6627 - Pampulha

Belo Horizonte - MG | Brasil

CEP 31270-901

revistapec@gmail.com

O CECIMIG agradece ao CNPq (Conselho Nacional de Desenvolvimento Científico) e à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) pela verba para a editoração deste artigo.